

Analyzing and Predicting Forestry Accountancy Network Variables with Bayesian Belief Networks as Compared to Traditional Analyzing Methods

Christoph Hartebrodt · Reinhard Aichholz ·
Marco Braasch

Accepted: 6 July 2010 / Published online: 29 July 2010
© Steve Harrison, John Herbohn 2010

Abstract The tremendous variability in physical conditions of forest enterprises as well as attitudinal aspects of their managers is seen as a major impediment to the understanding and optimization of forest management. For this reason, former studies using several methodological approaches—including meta analysis of econometric studies, binary choice models and stochastic frontier models—in many cases remained on a qualitative and more holistic level. This paper assesses the applicability of Bayesian Belief Networks (BBN) for the analysis of net income based on detailed 2006 economic data from the German federal accountancy network of forest enterprises larger than 200 ha. A network with one dependent (target) and 30 independent (explaining) variables was designed. The BBN has proven helpful for qualitative and to some extent quantitative analysis of economic data. It has become obvious that the completeness of populating the BBN model must be seen as a constraint. The speed of the calculations and the use of dependent probabilities can be seen as benefits of the BBN approach that reduce the risk of misinterpretation in comparison with traditional analysis methods such as the comparison of different strata. The visibility and presentability of the BBN approach facilitates its use in controlling and optimizing processes.

Keywords Systems analysis · Net income · Economic modeling · Sensitivity analysis · Forecasting

Introduction

One outstanding peculiarity of forest enterprises is their tremendous variability. This variability is not only confined to physical features such as size of the enterprise,

C. Hartebrodt (✉) · R. Aichholz · M. Braasch
Forest Research Institute Baden-Württemberg, Wonnhaldestraße 4, 79100 Freiburg, Germany
e-mail: christoph.hartebrodt@forst.bwl.de

tree species composition or topography, but also to the skills, attitudes and consequently behaviour of their owners or managers. These factors form a multi-causal framework that influences the forest activities and the financial, social and ecological outcomes of forest management. This variability must be seen as an impediment to the understanding and subsequently the optimization of forest management, because it is often difficult to separate outcomes resulting from management activities from outcomes due to external and non-controllable framework conditions. The ability to focus the analysis more closely on influenceable factors is crucial, because the effects of non-controllable factors are not an indication of inefficient management (Kuenzle 2005).

Past attempts to gain a deeper understanding of the key explanatory factors of forestry have in most cases been on a more 'holistic' level. Amacher et al. (2003) and Beach et al. (2005) undertook an extensive meta-analysis of econometric studies of forest management published during the last 25 years. It became apparent that the dependent variables are frequently qualitative or binary (e.g. harvest/no harvest, reforestation/no reforestation) and that the number of variables included usually does not exceed eight. Beach et al. (2005) and Amacher et al. (2003) assessed the methodologies applied and found that binary choice models including probit and logit functions are frequently used. Several studies relied on ordinary least square regression.

Some experience has been gained with quantitative methods as well. Hoffmann (2006) examined the applicability of data envelopment analysis in the comparative appraisal of the economic efficiency of forest enterprises. Lien et al. (2007) assessed the technical efficiency of harvesting using a stochastic frontier production function. They found that age and education had negative impacts on efficiency. Debt level, the existence of a management plan and the location of the enterprise in a rural region influenced the technical efficiency positively, as did combining forestry with agricultural activities.

With regard to identifying the most important aspects of forest management, Hartebrodt et al. (2005) found some interdependency between the felled volume and the financial operating result. However, the correlation coefficient was only 0.46, indicating that various other factors affect the operating result.

Insofar as the capability to explain the relevance of individual cost factors in forest enterprises remains rather low, there is still a need to identify and test new approaches to explain forest activities from the point of view of managerial economics. This paper examines the applicability of the Bayesian Belief Network (BBN) methodology to the economic data of forest management, particularly regarding the use of data from accountancy networks (ACNs) as inputs to modeling. The following research questions are investigated:

- Are BBN suited in general to analyze, explain and forecast economic data from forest accountancy networks?
- What are the strengths, weaknesses, opportunities and threats of the combination of BBN and ACNs?
- What are the important differences between BBN analysis and traditional forms of economic analysis of ACN data such as scenario analysis, comparison of substrata and regression analysis?

Besides these methodological aspects, an objective has been to test whether BBN methodology can be used to derive new, or to quantify existing qualitative, findings from earlier studies. The next sections introduce systems analysis based on the BBN methodology and provide an overview of data analysis using BBN. Details are then provided on the dataset applied in the study. Results of the BBN modeling are presented, followed by a SWOT analysis of the application of this methodology. Concluding comments follow, pointing to further prospects for the methodology.

Systems Analysis with Bayesian Belief Networks

According to Jackson (2003), cited in Smith et al. (2007), systems designed by human beings are ‘purposeful’ systems. The ‘system’ of a forest enterprise is, amongst other purposes, designed to achieve a positive financial outcome (Hartebrodt and Bitz 2007), particularly in the case of privately owned forest enterprises. However, the importance of the economic sphere varies considerably (Hartebrodt and Bitz 2007). One important approach for improving the management and consequently the operating result is to learn from the effects of previous management activities. Bosch et al. (2003) placed this in the context of adaptive management and offered a generalized adaptive management cycle (Fig. 1).

In recent years, BBN have gained importance in various sectors and are used in various decision-support systems. They are deemed to be useful instruments wherever incomplete information with a relevant level of uncertainty forms the basis for management decisions. McCann et al. (2006) provided an overview with regard to the useful characteristics of BBN models but also highlighted particular limitations. BBNs are used in medicine, electronic data processing and military sectors as well as for controlling production processes (Son 2002). A basic description of BBN can be found in Jensen (1996, 2001) and Nielsen and Jensen (2004). Basically BBN are models in which nodes (variables) are probabilistically related by some form of causal dependency. These causal dependencies are expressed by so-called ‘conditional probability tables’ (CPT; see Fig. 2). These

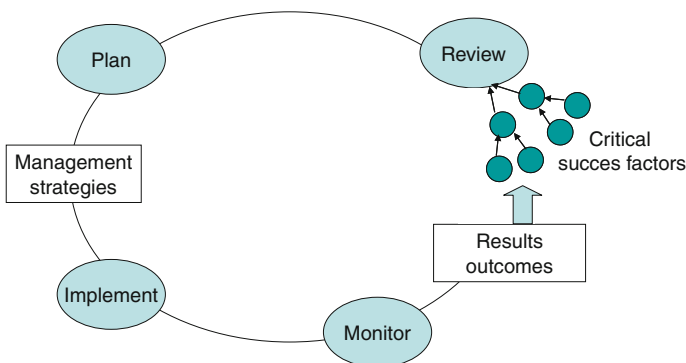


Fig. 1 Generalized adaptive management cycle. *Source:* Modified from Bosch et al. (2003)

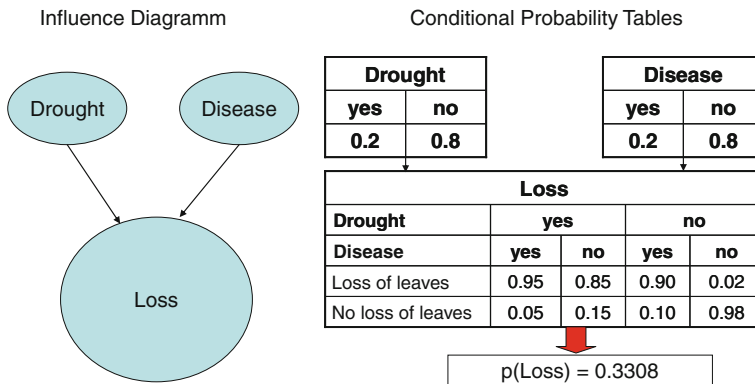


Fig. 2 BBN representing of the Apple Jack example with influence diagram and related conditional probability tables. *Source:* Adapted from Hugin (2010)

tables can be seen as the ‘heart’ of a BBN and hence it is important to inform these CPTs adequately.

An important feature of BBN is that various types of data sources can be used. Cain (2001) provided a systematization of four categories of data sources, namely (1) data obtained by direct measurement, (2) data obtained by stakeholder elicitation, (3) data from other models, and (4) data based on expert opinion. However, it should be noted that the explanatory power of the data decreases from category one to four.

For this study, the Netica[®] software package (Norsys Software Cooperation 1998) has been used. The database developed for the study contains category 1 data only. The influence diagrams have been converted into a BBN, which was ‘populated’ by extracting data from the ACN database directly using the BBN-software (see also Sect. “Design of a BBN using ACN Data”).

Data Analysis with Bayesian Belief Networks

Marcot et al. (2006) described the most important forms of data analysis with BBN. The basic methods of BBN-based analysis can be listed (*ibid.*) as follows:

- test of general behaviour of the model,
- test of sub-models,
- test of rank order of importance using sensitivity analysis.

These three forms are described next using one of the most famous BBN-examples, the ‘Apple Jack’ case (see also Fig. 2).

In this example, the domain is a small apple plantation belonging to Jack Fletcher (let’s call him Apple Jack). One day Apple Jack discovers that his finest apple tree is losing its leaves. Now he wants to know why this is happening. He knows that if the tree is dry (caused by a drought) there is no mystery—it is very common for trees to

lose their leaves during a drought. On the other hand the losing of leaves can be an indication of disease (Hugin 2010¹).

The situation can be modeled by the BBN in Fig. 2.² The BBN consists of three nodes: Disease, Drought, and Loss (of leaves) which can all be in one of two states: Disease can be either “yes” or “no”—Drought can be either “yes” or “no”—and Loss can be either “yes” or “no”. The node Disease tells us that the apple tree is sick by being in state “yes”. Otherwise, it will be in state “no”. The nodes Drought and Loss tell us in the same way if the tree is dry and if the tree is losing its leaves, respectively (Hugin 2010).

The *P*-value is the expected probability that the tree will lose its leaves. In the above example the calculation for *P*(Loss) would be:

$$\begin{aligned} P(\text{Loss}) &= 0.95 * 0.2 * 0.2 + 0.85 * 0.2 * 0.8 + 0.9 * 0.8 * 0.2 + 0.02 * 0.8 * 0.8 \\ &= 0.3308 \end{aligned}$$

It becomes clear that the total probability is the sum of the individual probabilities; $0.95 * 0.2 * 0.2$ is the probability of leaf loss if we have drought and disease (0.95) multiplied by the probability of drought (0.2) and disease (0.2).

Test of General Behaviour

It can be first tested whether the model generates expected outcomes. In this simple example the CPT can be used for this test. The highest probability of the loss of leaves is the combination of infection with a disease and drought, while the lowest probability of loss of leaves is the combination of no drought and no disease. The model meets expectations based on previous findings. In more complex models the *p*-values of loss related to different probabilities of individual states of explaining factors (e.g. yes = 0.3 and no = 0.7 for drought and disease) can be compared with the test of general behaviour, which is basically a comparison between the results generated by the model and previous empirical evidence.

Test of Sub-Models

The BBN analysis can be carried out at each point in the influence diagram. In the Apple Jack example it is only possible to base calculations at the probabilities of drought and disease, but if potential reasons for drought and disease were added (e.g. irrigation or climate, presence of pathogens or degree of resistance against infections) these sub-models (Drought and Disease) could be analyzed separately (see Fig. 3).

Test of Rank Order of Importance Using Sensitivity Analysis

According to Marcot et al. (2006, p. 3067), ‘Standard sensitivity analysis uses calculation of variance reduction when dealing with continuous variables and

¹ Cited from: <http://www.hugin.com/developer/getting-started/bns>.

² Here it is Fig. 2, in the cited online tutorial it is Fig. 1.

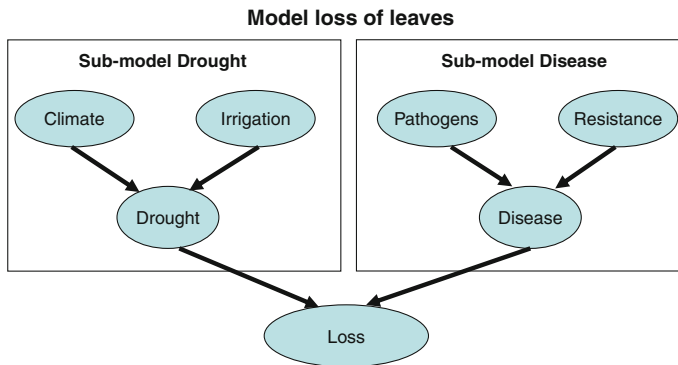


Fig. 3 Test of sub-models (extended Apple Jack example)

entropy reduction with discrete or categorical variables'. These authors introduced the mathematical background of variance and entropy reduction as well.³

A comparison of the sensitivities of the target variable related to individual parameters indicates the relative importance of the respective parameter in the model. Again the results of the sensitivity analysis obtained using the Netica[®]-tool have to be proven in the context of former empirical findings or expectations.

Accountancy Network Data of the Study Case

The completion of the CPTs of the highly explanatory accountancy network model based on reliable parameters is in general and in the present case an important step. Extensive data were available from a long-term economic monitoring system, the federal accountancy network (ACN) of forest enterprises. They have been used as results or outcomes according to the adaptive management cycle (as defined by Bosch et al. 2003; see Fig. 1) and build the base for deriving the underlying influence diagram. Following Smith et al. (2007) this influence diagram was, in a second step, converted into a system simulation model, in this case a Bayesian Belief Network.

The accountancy network contains data collected annually from forest enterprises with more than 200 ha of forested land. The ACN is designed for intensive economic documentation and is well suited for causative analysis. For this purpose about 700 individual pieces of data per enterprise are potentially available in the input layer of the database. The data set is structured into eight areas:

1. General information (spatial affiliation, kind of book-keeping, taxation...; number of attributes = 10)
2. Information about the forest area and owner (age, education,...; $N = 10$)
3. Description of stands (results of forest management planning, age class distribution, species; $N = 63$)

³ Formulae are presented in Sect. "Results of Sensitivity Analysis with BBN in Comparison with Regression Analysis".

4. Timber harvest and selling ($N = 100$)
5. Proceeds from various areas (timber, hunting, real estates,...; $N = 77$)
6. Structure of outlay and cost distribution sheet (staff, material, taxes,...; $N = 372$)
7. Staff (number, kind of employment,...; $N = 28$)
8. Other (membership of associations, share of unplanned felling, subsidies,...; $N = 36$)

For this preliminary research, only data from 2006 were used. Therefore there was no need to ‘normalize’ the data to avoid effects of changes in costs and prices (e.g. different increase rates of timber prices and input costs due to inflation). It is known that the quality of data differs substantially at the federal level due to varying standards of quality assurance between the individual states of Germany. Consequently, only data from Baden-Württemberg where quality assurance is executed in-house were used.

From the original data set with 700 pieces of data, a set of 31 key figures was derived for use in the study (Table 1). The selection was made with the support of a group of accounting experts from the state forest administration and scientists from within the Department of Forest Economics at the Forest Research Institute. The calculation rule for most of these 31 figures is stipulated and must be used. This rule is basically a hierarchical structure where the total revenues and total expenses are

Table 1 List of variables included in the BBN-model

<i>Target variable</i>		
Net income (€/ha)		
<i>Explaining variables, cardinal and continuous scale</i>		
Total expense (€/ha)	Total revenue (€/ha)	Revenue from subsidies (€/ha)
Timber sales revenue (€/ha)	Non timber revenues (€/ha)	Operating expense (€/ha)
Administrative expense (€/ha)	Annual felling (m^3/ha)	Proportion of beech sold (%)
Proportion of spruce sold (%)	Revenues from hunting (€/ha)	Revenues from property (lease) (€/ha)
Other revenues (€/ha)	Revenues from by-products (€/ha)	Expense for silvicultural treatment (general sense) (€/ha)
Expense for forest infrastructure (roads, dirt roads) (€/ha)	Expense for timber harvesting (€/ha)	Expense for reforestation (€/ha)
Felling budget (m^3/ha)	Expense for forest protection (€/ha)	Expense for hunting (€/ha)
Expense for precommercial thinning (€/ha)	Expense for by-products (€/ha)	Expense for maintenance of real estate (€/ha)
Expense for logging (€/ha)	Expense for skidding (€/ha)	Length of forest roads and tracks (m/ha)
Other expenses (€/ha)	Other expenses general sense (€/ha)	
<i>Explaining variables, nominal and discrete scale</i>		
Forest type (mainly conifers/mixed/mainly broadleaves)		

summed in a stepwise manner, beginning with detailed input values (e.g. logging expenses, see Fig. 5) to form subtotals for various areas (e.g. harvesting expenses, see Fig. 5). The calculation rule follows the logical structures of the outlay and proceeds and was used as a basic structure for the causal net or influence diagram presented in Fig. 5.

Design of a BBN using ACN Data

The original dataset was transferred from the underlying MS Access[®] relational database to a flat data matrix (spreadsheet) in MS Excel[®]. The Excel data could be integrated automatically into Netica[®], with individual variables displayed as boxes. The network has been built by linking the individual boxes graphically with arrows. The design of the network, predefined by the influence diagram and depicted in Fig. 4, took no more than 20 s per box. Consequently, the model could be built in less than 15 min, assuming that the user is familiar with the BBN concept and modeling procedures (including the Netica[®]-tool).

Because the BBN is based on the percentage of cases in each class, it is necessary to make the dataset discrete. This can be done manually or automatically, but even under the latter approach the number of states can be freely chosen. After the discretization of the data the model can be used to perform policy simulations.

Based on the existing rules for the calculation of key figures, stemming from an agreement between the federal ministry and the operators of the accountancy networks in the federal states, the network was designed as illustrated in Fig. 5. Net

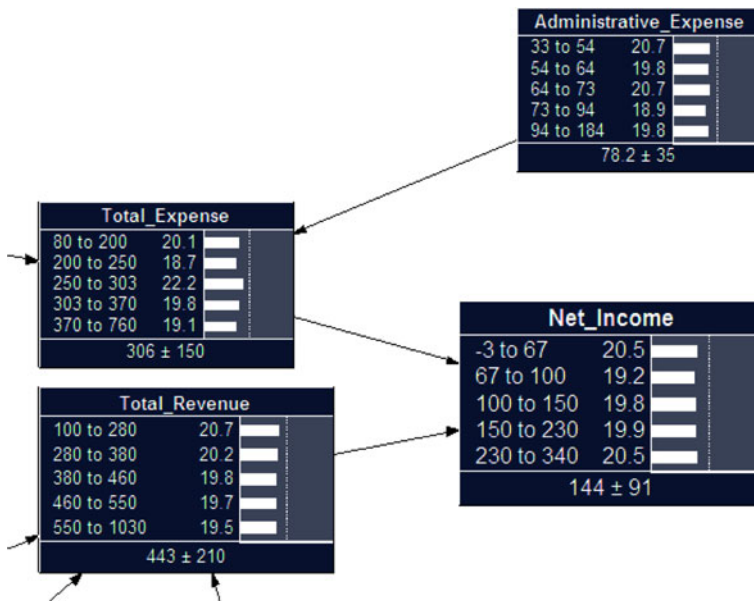


Fig. 4 Data boxes and arrows with 5 state-discretization of dataset

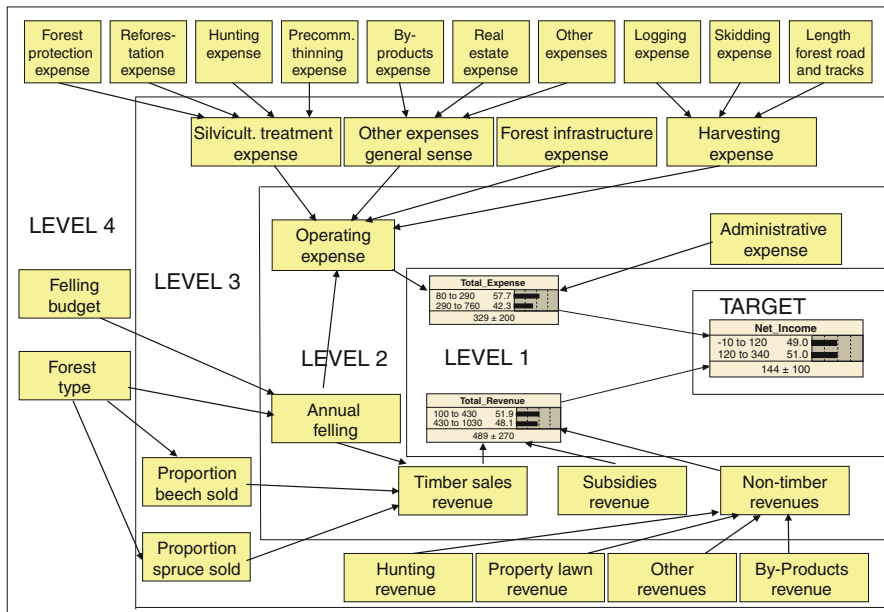


Fig. 5 The Bayesian Belief Network for explaining net income of a forest enterprise. For reasons of readability, only 'target' and 'Level 1' boxes are displayed as Netica boxes

income is used as the target (dependent) variable, which is expected to be influenced by 30 explanatory (independent) variables. Despite the absence of a strong hierarchical structure, it is apparent that the variables are grouped into four levels that are characterized by an increasing 'distance' to the target variable. Basically, variables that are related to expenses are placed in the upper part of the figure and those that are related to revenues in the lower part. But not all variables are measured in financial units. A strength of BBN is that non-financial explanatory variables can be integrated into the model. Factors relying on expert opinion such as felling budget, annual felling amount, forest type and other data including proportions of spruce or beech sold have also been integrated. They can be found in the left side of levels three and four of the net structure of the explaining variables (as shown in Fig. 5).

The means and standard deviations of cardinal variables are presented in the lower part of the individual box (e.g. in Fig. 4 and partially in Fig. 5).

Applicability of the BBN Method for Analysis, Explanation and Forecast of Accountancy Network Data

ACN data are frequently used for the calculation of averages, performing sensitivity analysis, forecasting and the transfer of general knowledge to individual enterprises. The differences between traditional approaches and BBN-based approach are now examined.

Table 2 Comparison of means with different discretizations in ACN-BBN model

Variable	Arithmetic mean, individual values (=100%)	BBN discretization 2 state		BBN discretization 5 state	
		Value of arithm. mean/median	Value Difference (%)	Value Difference (%)	
Net income (€/ha)	141/121		144 2.35	144 2.35	
Total expense (€/ha)	286/284		329 14.85	306 6.82	
Operating expense (€/ha)	212/211		267 26.15	232 9.61	
Administrative expense (€/ha)	74.8/69.6		88.6 18.43	78.2 4.53	
Total revenue (€/ha)	427/429		489 14.47	443 3.71	
Revenue from timber (€/ha)	399/401		460 15.28	419 5.00	
Revenue from subsidies (€/ha)	4.8/0.0		16.1 237.00	6.45 35.01	
Revenues from non-timber-products (€/ha)	23.3/17.6		39.0 67.04	28.2 20.78	

Calculation of Averages

Related to the probabilistic structure of the BBN, the calculation of means is based on the frequencies of the individual states (of one variable) multiplied by the means of the respective state. It is of interest to examine whether this approach leads to values that meet the arithmetic mean or the median calculated from the individual values. It can be shown that the similarity of the means (BBN-classes and individual values) is related to the number of states in the BBN-model. When using the arithmetic mean based on individual values as a standard, discrepancies can be found between the BBN findings and the calculation based on individual values. On the basis of automatic two- and five-state discretization, differences between 2.35 and 237% are obtained. Especially when lower absolute values are included, the probability of making an error when calculating means in BBN-models increases. In addition, consistent with expectations, the findings suggest that better results, in terms of a smaller difference between arithmetic mean and values achieved using BBN models, are obtained when a five-state discretization is used. In this case, the discrepancy has a maximum of about 35% and in most cases is below 10% (Table 2).

Results of Sensitivity Analysis with BBN in Comparison with Regression Analysis

The Netica[®] tool provides an automatic sensitivity analysis based on the reduction of the variance or entropy of data. It can be shown that only a few variables have a meaningful impact on the net income. Percentage reductions of the variance of 0.091⁴ and 0.054 for total revenue and timber sales revenue respectively, highlight the most important factors.

⁴ According to Marcot et al. (2006, p. 3074), ‘The following formulae are used to calculate model sensitivity in the modelling shell Netica[®] (B. Boerlage, personal communication). Variance reduction

A traditional method for analyzing the interdependencies between variables is a regression analysis. The results of the 18 most important variables (in terms of the sensitivity of net income to the respective variable) depicted in Table 3 show a considerable overlap between the R^2 values and the reduction of the percentage sensitivity values provided by the BBN model. Other studies have supported the hypothesis that the sensitivity of performance variable in BBN models is at least partially related to the position in the net. However, the findings in this case only partially support this assumption. A higher sensitivity of the target variable to five variables which are located in the third and fourth levels is detected than for all the variables in the second level. In the present case the results indicate that a variable that shows a higher R^2 (related to the target variable) causes a higher sensitivity too, relatively independent of its position in the net (Table 3). Annual felling and felling budget are definitely explaining factors for operating expenses and timber sales revenue. This must be taken as an indication of their importance and not as a need to redesign the BBN, in terms of moving these variables into another level. Together with the results of the regression analysis, the high sensitivity of net income related to these variables confirms their importance and reveals that this is at least partially independent of their position in the net.

Prediction with BBN Using a Single Variable and Based on Modification of Individual Input or Output Factors

The Netica[®] tool provides various ways to predict the target variable. It is possible to vary the likelihood of each state of every explaining variable. In this case (and with the aim of comparison with traditional analyses) the potential net income was analyzed with Excel[®] and Netica[®].

The prognostication of economic results at the enterprise level is considered difficult by most forest economists, especially as multiple explanatory components always influence forest entrepreneurial success. It is scarcely known how forest enterprises act because this is dependent on physical and natural conditions, and no two forest enterprises are identical. Furthermore, at least until recently, little has been known about the functional relationships between explaining and explained factors and little was known about the interdependencies between explaining factors. Therefore approaches that predict the net income—e.g. by the modification of individual variables while leaving other variables uninfluenced—are not expected to produce informative findings. One frequently applied approach is the selection of

Footnote 4 continued

(VR) is the expected reduction in the variation, $V(Q)$, of the expected real value of the output variable Q having q states due to the value of an input variable F having f states, and is calculated as $VR = V(Q) - V(Q|F)$, where $V(Q) = \sum_q P(q) [Xq - E(Q)]^2$, $V(Q|F) = \sum_q P(q|f) [Xq - E(Q|f)]^2$, $E(Q) = \sum_q P(q) Xq$, where Xq is the numeric real value corresponding to state q , $E(Q)$ is the expected real value of Q before any new findings, $E(Q|f)$ is the expected real value of Q after new findings f for mode F , and $V(Q)$ is the variance of the real value of Q before any new findings. Entropy reduction, I , is the expected reduction in mutual information of Q (measured in information bits) due to a finding at F , and is calculated as $I = H(Q) - H(Q|F) = \sum_q \sum_f \frac{P(q,f) \log_2 [P(q,f)]}{P(q)P(f)}$ where $H(Q)$ is the entropy of Q before any new findings and $H(Q|F)$ is the entropy of Q after new findings from node F' .

Table 3 Comparison of BBN sensitivities and coefficient of determination

Variable	Percentage variance reduction (sensitivity of revenue)	Position in the BBN net (level)	R^2 adj
Net income	0.2498931 ^a		
Total revenue	0.0912561	1	0.60
Timber sales revenue	0.0538631	1	0.59
Annual felling	0.0191046	3	0.47
Total expense	0.0077791	1	0.08
Forest type	0.0047777	4	0.11
Proportion of spruce sold	0.0042748	3	0.15
Felling budget	0.0038852	4	0.21
Proportion of beech sold	0.0032313	3	0.11
Operating expense	0.0016783	2	0.06
Administrative expense	0.0010916	2	0.05
Revenue from subsidies	0.0007850	2	0.06
Expense for timber harvesting	0.0005435	3	0.08
Other expenses	0.0001512	3	<0.01
Expense for logging	0.0001314	4	0.06
Expense for silvicultural treatment	0.0001204	3	<0.01
Expense for forest infrastructure	0.0001203	3	<0.01

^a The percentage variance of the target variable is an indicator how much the degree of randomness can be reduced in total

sub-strata, which are defined by cut-offs of single or multiple indicators, where the averages of the target variable of these sub-strata are compared. This can be carried out using both the table function of Excel[®] and the dependent probability function of a BBN. In the latter case, modifications are made by using a 100% likelihood of individual states. Table 4 and the example in Fig. 6 present a comparison of percentage differences in net income between the average net income for the best or highest (max) and worst or lowest (min) part of the population, selected with regard to various explanatory variables. The column heading ‘Excel 2’ (substrata) and ‘Netica 2’ (substrata) indicate that the population is divided into two substrata. ‘Excel 5’ and ‘Netica 5’ denote that the substrata are composed by the enterprises representing 20% of the best or highest and worst or lowest values of the individual explaining variable.

Figure 6 depicts the two approaches using an example dataset. In the Excel[®] two-state approach the two means of the operating result of the substrata, defined by the upper (197) and lower (88) half of enterprises according to the amount of annual felling, are compared and lead to a percentage between the minimum (min) and maximum (max) values of 124%. Using the Netica[®] approach, the likelihoods of states 5.6–8.4 and 8.4–10.2 are changed to 100%. This results in calculated net incomes of 177 and 128 and a percentage change of 36% in net income. Because of the use of an example dataset there are some slight differences between the mean values of the Netica 2 state calculation in the box in Table 4 and Fig. 6.

Table 4 Comparison of average net income of best and worst sub-strata defined by explanatory variables and five and two-class discretization

Explanatory variable	Change expected (%)	Excel 2	Change max to min (%)	Netica 2	Change max to min (%)	Excel 5	Change max to min (%)	Netica 5	Change max to min (%)
Total revenue (€/ha)	Min Max	84 198	136	93 199	114	48 231	381	116 180	55
Total expense (€/ha)	Min Max	123 162	32	158 126	-20	93 167	80	134 147	10
Revenue timber (€/ha)	Min Max	85 196	131	105 186	77	51 242	375	139 149	7
Revenue subsidies (€/ha)	Min Max	113 170	50	139 149	7	113 199	76	143 145	1
Non-timber revenue	Min Max	147 145	-1	145 144	-1	147 145	-1	145 144	-1
Operating expense (€/ha)	Min Max	120 161	34	151 137	-9	83 161	94	140 145	4
Administrative expense (€/ha)	Min Max	127 155	22	150 138	-8	124 194	56	142 145	2
Annual felling (m ³ /ha)	Min Max	88 197	124	121 170	40	48 226	371	143 145	1

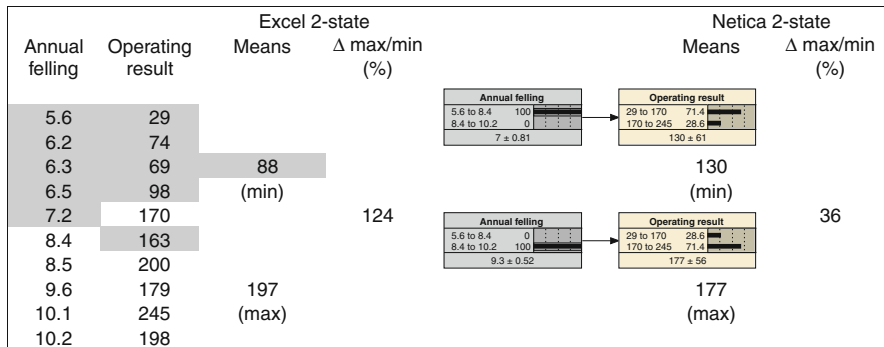


Fig. 6 Demonstration of calculation of percentage differences between Excel[®] and Netica[®]/BBN based calculation

It is obvious that sensitivities depend widely on the method of discretization and selection. The five-sub-collective Excel[®] approach shows, in general, the highest sensitivities (here expressed as percentage $\Delta \text{ max/min}$), followed by the two-sub-collective Excel[®], the two-state Netica[®] model and then the five-state Netica[®] model. In the latter, only total revenue and total expense exhibit any meaningful sensitivity in the results. The Excel[®] modeling, which only varies the respective factor (variable), shows stronger effects than when using the Netica[®] approach with the same discretization. Hence, there is a notable leveling effect due to the state-based calculation of averages. Two-class models in Excel[®] show approximately equal change rates in comparison to the two-state Netica[®] model, but in some cases opposite signs. In both five-class models, there are large differences in the percentage change rates, mainly as a result of the leveling effect.

When the expected direction and strength of change is compared with the prognosed results, a qualitative overlap exists in most cases between expectation and modeling. However, this is not the case for the Excel[®] results with regard to the expense-related variables. Here an increase in expenses does increase net income, which at first glance is not rational. In general it can be stated that the best qualitative overlap between the expected and calculated change rate takes place in the Netica[®] two-state model.

Transferability to Individual Enterprises

With regard to the question of whether the BBN model can be used to predict the net income for the whole range of forest enterprises, it is of interest to see what percentage of the observed values are covered by results from modeling. The results for those continuous variables causing relevant sensitivities in the target variable (operating result) are depicted in Table 5. Here the share of observations that are below or above the minimum and maximum value calculated by the Netica[®] model is reported. Because only the two-state Netica[®] model revealed significant sensitivities between the upper and lower level, only this model is depicted in Table 5.

Table 5 Coverage of individual enterprise values by the two-state Netica® model

Variable	Difference arithm. mean/median (%)	Enterprises below threshold (%)	Enterprises in 'Netica' range (%)	Enterprises above threshold (%)
Net income	16.4	23.6	51.9	24.5
Annual felling	0.9	7.5	91.5	0.9
Total expense	0.6	0.9	94.3	4.7
Operational expense	0.5	8.5	90.6	0.9
Administrative expense	7.5	17.0	77.4	5.6
Total proceeds	−0.6	9.4	86.8	3.8
Proceeds from timber	−0.4	15.1	89.0	1.9
Subsidies	Not applicable	0	97.2	2.8
Proceeds from non- timber products	32.4	20.7	76.4	2.8

Speed of Developing BBN Models and Data-Processing

Assuming that the dataset is available and the structure of the net is predefined as in the present case, the time required to design and finally run the BBN model is less than 30 min. The sensitivity analysis can be executed in a few minutes; each prognostication can be done with one or two clicks of the mouse.

Changes in the network structure can be executed easily by introducing new or deleting existing arrows between the nodes (variables). Thus, exploratory data mining can be performed easily. The technical process takes only a few minutes with basic analyses such as calculation of arithmetic means, frequency distributions and class limits of uniformly populated classes being an implicit part of the network design. Consequently, these steps do not require additional computing time. In comparison with other approaches, the BBN method provides a much faster way to perform data mining on data with a complex structure. It is more flexible than other methods that normally require individual calculations to be performed.

Completion of the Causal Probability Tables

BBN are basically able to run with incomplete data. However, the results from the first output level provide no information as to whether all results of BBN analysis are based on information or not. The amount of data processed in each case (a case is defined as an individual probability in the CPT) can be taken as a measure of the 'degree' of information. Table 6 provides information about this degree for a selection of CPTs. The share of individual cases that are not informed, or are only weakly informed, varies considerably. This is due to the fact that the number of cases is the number of states raised to the power of the number of parent nodes plus one. For the study reported here, in which 106 datasets from the ACN results of 1 year are used, only the two-state version can be considered sufficiently informed.

Table 6 Degree of information in various CPTs and different discretizations

Variable or node	Number of parent nodes	Number of states	Number of cases	Number of cases without data	1–10 datasets per case	>10 datasets per case
Net income	2	2/5	8/125	2/89	0/36	6/0
Total expense	2	2/5	8/125	2/87	3/38	3/0
Total revenue	3	2/5	16/625	4/551	6/74	6/0
Operating expense	5	2/5	64/15625	33/15524	94/101	1/0
Timber revenue	3	2/5	16/625	4/551	10/74	2/0
Non-timber revenue	4	2/5	32/3125	5/3034	26/91	1/0
Annual felling ^a	2	2 (3)/5 (3)	8/75	0/31	3/43	5/1
Non-timber revenue	4	2/5	16/625	1/534	10/91	5/0

^a One parent note has three states

SWOT Analysis of the Accounting Network BBN

The approach of identifying the strengths (to be utilized), weaknesses (to be overcome or minimized), opportunities (to be exploited) and threats (to be avoided) has found wide use as an analysis framework in relation to particular industry and policy initiatives (Schmitthüsen et al. 2003). The analysis is typically carried out subjectively in group discussions involving people with particular knowledge and experience in the subject area. In the present case this SWOT-analysis was carried out with a small group of ACN experts.

Strengths of BBN Modeling

Empirical evidence suggests that there is only limited knowledge about the nature of interrelationships between various influencing factors. Despite the availability of detailed knowledge on the present state of individual variables, further attempts to derive explanatory models for financial performance variables remained incomplete or qualitative. BBN-based sensitivity analyses can be seen as a promising approach to visualize and quantify these basic interdependencies. The BBN diagram visualizes the interdependencies and is a suitable tool to discuss economic results with both practitioners and scientists.

It can be shown that there is a tendency toward cost-stability and decreasing cost-variability between individual enterprises when examining the financial performance estimates of the ACN (Fillbrandt and Hartebrodt 2006; Hartebrodt et al. 2007). It

follows that the marginal benefits of cost reduction should be low,⁵ and the results depicted above support this view. Both regression analysis and the BBN-model approach reinforce this awareness, despite the fact that many forest enterprises still focus on cost reduction. The BBN-based analysis clearly identifies this causal relationship. In addition, even the underlying influencing factors including forest type and annual harvest are identified correctly. To this extent, BBN models can be seen as a convenient tool to analyze and depict, at least qualitatively, relationships in complex economic structures. The test of general behaviour of the model proved successful.

The visible network structure can itself be an issue of discussion between experts and avoids the impression of a black box system. As economic optimizations are under discussion in most cases, this open structure can be seen as an essential part of acceptance management when introducing the need for change to managers and employees by upper management.

Analyzing and comparing the ability to prognose the impact of alternative management activities—e.g. increasing the annual cut or reducing costs—shows some of the advantages of the BBN approach. The risk of misinterpretation is reduced because BBN models always rely on all data from the entire network. As a consequence, the most relevant finding is the influence of cost reduction or increase in net income from forestry. Both Excel[®] models indicate that an increase in costs would increase net income. In contrast, the Netica[®] two-state model predicts a decrease in the operating result. The reason for this difference is that Netica[®] is able to consider the relationship between the annual cut and costs. In the Excel[®] models, the identification of various sub-populations on the basis of costs leads implicitly to two different substrata with regard to the annual cut, which is known to be positively correlated with net income. These implications cover the effects due to increasing cost and lead to misleading results. In this regard, assuming the appropriate discretization is used and there is enough information to populate the model properly, the BBN approach reduces the risk of misinterpretation.

Weaknesses of BBN Modeling

In that arithmetic means can be misleading, for example in the case of a large number of outliers or asymmetric distributions, the discrepancies of means in BBN in comparison with the calculation based on individual data must be seen as an impediment for the use of BBN. These faults are primarily related to the state-frequency-based calculation and interferences between nodes. The findings suggest a considerable risk of misinterpretation when using the means of cardinal-scaled variables. The risk decreases in terms of the relative deviance with higher absolute values, a low standard deviation and, more importantly, when the number of states can be increased, due to the availability of a higher number of data points.

The leveling effect induced by the state frequency calculation of the means remains a constraint. Table 5 reveals that the prognostication ability for individual

⁵ The low cost variability leads directly to a low variance and therefore to a low sensitivity in the BBN (Netica[®]) model. In the substrata model the low variability causes a small difference between the averages of the highest and lowest substrata.

enterprises is limited because of the varying proportion of real values that is outside the range that can be prognosticated by BBN models. This seems to be particularly the case whenever asymmetric frequency distributions or a large number of outliers are present.

Opportunities of BBN Modeling

The use of data from more than one financial year can be seen as a promising way to inform the CPTs more completely. In that BBN models cannot be used for prognoses when individual enterprises have extreme values for one or more key figures, using data from more than 1 year can be expected to reduce the number of outliers considerably. However, it has to be considered that in this case the data would have to be normalized to level effects of inflation and price change, for instance. The use of averages of different years is an option for further improvement, but does not lead to a better completion of the CPTs because the number of datasets remains stable.

Another promising approach is the stepwise analysis of BBN. Despite the evidence that position in the net only partially explains the sensitivity of individual factors on financial performance, it is apparent that factors that are ‘parents’ to influencing factors that cause a low sensitivity of the target variable also cause low sensitivity of the target variable. However, it is possible to analyze the sensitivities of the dependent variable related to each position in the net. Wood production, for example, is an explanatory factor for net income but is of low importance (sensitivity value of net income = 0.00012; 13th rank; see Table 3). The related factors—forest protection, reforestation, hunting expense and silvicultural treatment—rate much lower in terms of the sensitivity values of the target variable (between 0.00002 and 0.000005). However, it is possible to run the sensitivity analysis for cost of wood production as well. The sub-net structure (Fig. 7) and sensitivity-values (Table 7) are provided below.

Table 7 reveals that costs of reforestation influence the costs of wood production to the greatest extent in comparison with the other factors listed above. One can consecutively

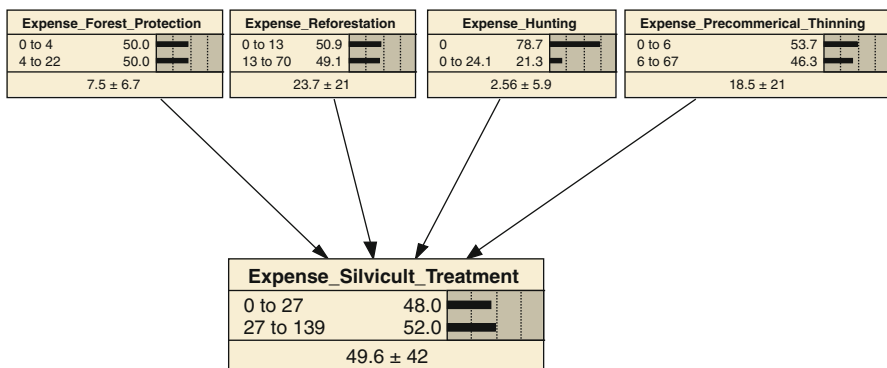
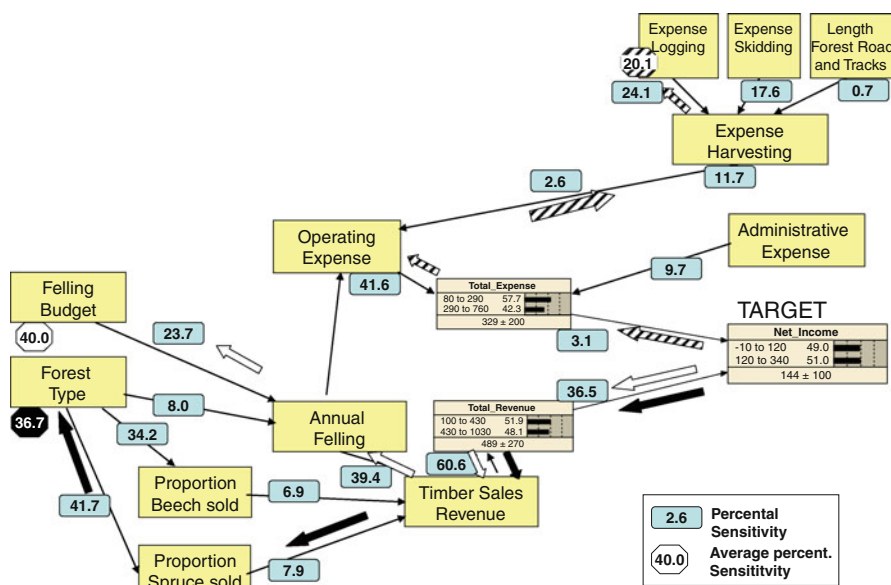


Fig. 7 Excerpt from the BBN net for analysis of key influence factors on costs of wood production

Table 7 Influencing factors on the cost of wood production

Variable	Variance reduction (production cost sensitivity)	Variance reduction (%)
Silvicultural expenditure Silvicultural treatment (target)	0.249	100.0
Reforestation expense	0.046	18.5
Forest protection expense	0.017	6.9
Pre-commercial thinning expense	0.003	1.6
Hunting expense	0.001	0.4

**Fig. 8** Causal chains in the ACN net for the explanation of net income

use the sensitivity analysis of sub-models to identify causal chains or to reduce the complexity of a BBN. Figure 8 depicts the main causal chains in the ACN network.

The average of the sensitivities can be used as an indicator of the strength of the causal chain. From Fig. 8 it is apparent that felling budget and forest type (classified as mainly spruce, mainly beech or predominately mixed stands) are the key explaining components. From this it follows that the analysis of sub-models provides valuable information as well as the analysis of the whole model.

Threats of BBN Modeling

BBN provide a rapid insight into the nature of interrelationships between individual explanatory factors. However, results generated by BBN models have to be interpreted carefully. It is necessary to carry out a critical review of how intensively and uniformly

populated the individual node is. This is especially the case when the automatic interface of Netica[®] is used for data input and definition of CPTs. In combination with the formation of more than two or three states, partially informed nodes strongly influence the sensitivity of the BBN model and the calculation of means. Averages of continuous variables are highly sensitive to the number of states. This can be explained by the state- and frequency-driven calculation of the means. Highly discretized models, which contain cardinal values, tend to have a strong leveling effect on averages, due to the fact that the changes in the frequency distribution of the individual state are moderate and consequently the averages tend to be stable.

There is a trade-off between the sensitivity of BBN models and their mathematical correctness. The more states used, the more precise the calculation of the means will be. These means are important because they provide valuable information on expected costs and revenues in various sub-populations or economic framework conditions (e.g. type of forests that cannot be modified over a short timescale). However, this mathematical ‘correctness’ reduces the ability of the BBN model to predict absolute changes in cardinally measured values. In using 106 datasets only two or three state BBN models provided satisfactory results.

Concluding Comments and Proposed Developments

BBN modeling can be seen as a helpful tool for the analysis of ACN data. Given that the ACN datasets are available, little time is required for the development of the network, basic sensitivity analysis and the calculation of means. BBN models provide a qualitative overview of relationships between individual variables. Their inherent ability to simulate forest management decisions is better suited to informing forest practitioners than many other tools. The possibility to modify each variable in the net and estimate the reactions of related factors helps in the understanding of the underlying financial relationships between components of these complex economic structures. To this extent, the derivation of basic concepts for the optimization of forest enterprises is possible. The sensitivity analysis has also proven to be a helpful tool. Therefore BBN models can be used by scientists who want to gain initial and quick insights into detailed datasets. For practitioners, especially controllers, who want to explain basic concepts of entrepreneurial improvement, the visualization of the results can be considered helpful. The ability to run sensitivity analyses and to some extent show the possible effects of management activities can ease the communication between controller and managers.

BBN models based on ACN datasets can also be used for the prognostication of absolute values of costs and revenues for individual enterprises. However, for all enterprises that show particular peculiarities in important economic key factors (those inducing high sensitivities of the explained variables), prognostication would be misleading. For other enterprises the leveling effect of BBN models may underestimate change rates.

The BBN approach is a helpful methodology in explaining financial results of forest enterprises. Its speed of data mining is a considerable advantage compared to

other analytical tools. BBN modeling can be seen as a helpful supplement to the analytical toolbox for forest economists, but not as a silver bullet applicable without critical review.

References

- Amacher GS, Conway MC, Sullivan J (2003) Econometric analyses of nonindustrial landowners: is there anything left to study. *J For Econ* 2003(9):137–164
- Beach RH, Pattanayak SK, Yang J-C, Murray BC, Abt RC (2005) Econometric studies of non industrial private forest management a review and synthesis. *For Policy Econ* 7(2005):261–281
- Bosch OJH, Ross AH, Beeton RJS (2003) Integrating science and management through collaborative learning and better information management. *Syst Res Behav Sci* 20(2):107–118
- Cain J (2001) Planning improvements in natural resources management: guidelines for using Bayesian Networks to support the planning and management of development programs in the water sector and beyond. Centre for Ecology and Hydrology, Wallingford UK
- Fillbrandt T, Hartebrodt C (2006) Forstwirtschaft lohnt sich wieder. *AFZ-DerWald* 23(2006):1246–1248
- Hartebrodt C, Bitz S (2007) From framework to forest activities. *Small Scale For* 6(3):309–328
- Hartebrodt C, Fillbrandt T, Brandl H (2005) Community forests in Baden-Württemberg (Germany): a case study for successful public-public-partnership. *Small Scale For* 4(3):229–250
- Hartebrodt C, Fillbrandt T, Hercher W (2007) Einblicke in Entwicklungen im baden-württembergischen Waldbesitz -Von schwarzen Zahlen und Aufwandskonstanz. *AFZ-DerWald* 22(2007):1185–1187
- Hoffmann C (2006) Die Data Envelopment Analysis (DEA) und ihre Anwendungsmöglichkeiten zur vergleichenden Effizienzanalyse im Forstwesen. Doctoral Thesis, University of Natural Resources and Applied Life Sciences, Vienna
- Hugin (2010) Hugin BBN online tutorial. <http://www.hugin.com/developer/getting-started/bns>. Accessed 15 June 2010
- Jackson MC (2003) Systems thinking: creative holism for managers. Wiley, Chichester
- Jensen FV (1996) An introduction to Bayesian networks. Springer, New York
- Jensen FV (2001) Bayesian networks and decision graphs. Springer, New York
- Kuenzle M (2005) Cost efficiency in network industries: application of stochastic Frontier analysis. Dissertation Swiss Federal Institute of Technology, Zürich
- Lien G, Stordal S, Baardsen S (2007) Technical efficiency in timber production and effects of other income sources. *Small Scale For* 6(2):65–78
- Marcot B, Steventon JD, Sutherland GD, McCann RK (2006) Guidelines for developing and updating Bayesian belief networks applied to ecological modeling and conservation. *Can J For Res* 36(12):3063–3074
- McCann RK, Marcot BG, Ellis R (2006) Bayesian belief networks: applications in ecology and natural resource management. *Can J For Res* 36(12):3053–3062
- Nielsen TD, Jensen FV (2004) Learning a decision maker's utility function from (possibly) inconsistent behaviour. *Artif Intell* 160(1/2):53–78
- Norsys (Software Cooperation) (1998) Netica application user's guide. Norsys Software Corp. Vancouver, BC, Canada. http://www.norsys.com/tutorials/netica/nt_toc_A.htm. Accessed 20 Jan 2009
- Schmittthüsen F, Kaiser B, Schmidhauser A, Mellinghoff S, Kammerhofer AW (2003) Unternehmerisches Handeln in der Wald- und Holzwirtschaft. Deutscher Betriebswirte-Verlag GmbH, Gernsbach
- Smith C, Felderhof L, Bosch OJH (2007) Adaptive management: making it happen through participatory systems analysis. *Syst Res Behav Sci* 24(6):567–587
- Son S (2002) Einsatz von Bayesschen Netzwerken zur Simulation von Unternehmensentscheidungen. Seminararbeit am Lehrstuhl für Betriebswirtschaftslehre der Johann Wolfgang Goethe Universität Frankfurt